Optimal Transmission Strategy and Capacity Region for Broadcast Z Channels

Bike Xie, Miguel Griot, Andres I. Vila Casado and Richard D. Wesel
Department of Electrical Engineering, University of California, Los Angeles, CA 90095-1594
Email: xbk@ee.ucla.edu, mgriot@ee.ucla.edu, avila@ee.ucla.edu, wesel@ee.ucla.edu

Abstract— **This paper presents an optimal transmission strategy, with simple encoding and decoding, for the two-user broadcast Z channel. This paper provides an explicit-form expression for the capacity region and proves that the optimal surface can be achieved by independent encoding. Specifically, the information messages corresponding to each user are encoded independently and the OR of these two streams is transmitted. Nonlinear turbo codes that provide a controlled distribution of ones and zeros are used to demonstrate a low-complexity scheme that works close to the optimal surface.**

Index Terms— **broadcast channel, broadcast Z channel, capacity region, turbo codes.**

I. Introduction

Degraded-broadcast channels were first studied by Cover in [1] and a formulation for the capacity region was established in [2], [3] and [4]. The key idea to achieve the optimal surface of the capacity region for degraded-broadcast channels is superposition coding. In superposition coding for degraded-broadcast channels, the data sent to the user with the most degraded channel is encoded first. Given the encoded bits for that user, an appropriate codebook for the second most degraded channel user is selected, and so forth. Hence superposition coding is in general a joint encoding scheme. However, combining independently encoded streams of each user is an optimal scheme for some broadcast channels including broadcast Gaussian channels [1] and broadcast binary-symmetric channels[1] [2].

This paper focuses on the study of broadcast Z channels. The Z channel is the binary-asymmetric channel shown in Fig. 1(a). Fig. 1(b) shows a two-user broadcast Z channel. This paper provides an explicit-form expression for the capacity region of the two-user broadcast Z channel and proves that independent encoding with successive decoding can achieve this capacity region.

This paper is organized as follows. In Section II, some definitions and notation for broadcast channels are introduced. The proof that independent encoding can achieve the optimal surface of the capacity region for the two-user broadcast Z channel is presented in Section III. Nonlinear-turbo codes, designed to achieve the optimal surface, are presented in Section IV and simulation results are shown in Section V. Section VI delivers the conclusions.

This work was supported by the Defence Advanced Research Project Agency SPAWAR Systems Center, San Diego, California under Grant N66001-02-1-8938 and by the state of California and ST Microelectronics through UC Discovery Grant 03-10142.

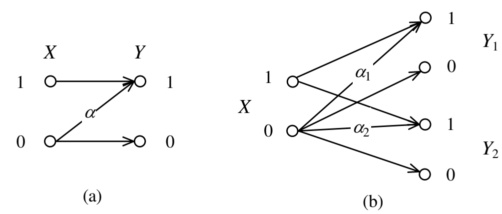

Fig. 1. (a) Z channel. (b) Broadcast Z channel.

II. Definitions and Preliminaries

A. Degraded broadcast channels

The general representation of a discrete memoryless broadcast channel is given in Fig. 2. A single signal X is broadcast to M users through M different channels. Channel A_2 is a physically degraded version of channel A_1 and broadcast channel $X \rightarrow Y_1, Y_2$ is physically degraded if $p(y_1, y_2|x) = p(y_1|x)p(y_2|y_1)$ [5]. A physically degraded broadcast channel with M users is shown in Fig. 3. Since each user decodes its received signal without collaboration, we only need to consider the marginal transition probabilities $p(y_1|x), p(y_2|x), \cdots, p(y_M|x)$ of the component channels $A_1, A_2, \cdots, A_M$. Since only the marginal distributions affect receiver performance, a weaker notion of a stochastically degraded broadcast is defined in [2] and [5].

Let A_1 and A_2 be two channels with same input alphabet $\mathcal{X}$, output alphabet $\mathcal{Y}_1$ and $\mathcal{Y}_2$, and transition probability $p_1(y_1|x)$ and $p_2(y_2|x)$ respectively. A_2 is a degraded version of A_1 if there exists a transition probability $q(y_2|y_1)$ such that

$$p_2(y_2|x) = \sum_{y_1 \in \mathcal{Y}_1} q(y_2|y_1)p_1(y_1|x).$$

A broadcast channel with receivers $Y_1, Y_2 \cdots, Y_M$ is a stochastically degraded broadcast channel if every component channel A_i is a degraded version of A_{i-1} for all $i = 2, \cdots, M$ [2]. Since the marginal transition probabilities $p(y_1|x), p(y_2|x), \cdots, p(y_M|x)$ completely determine a stochastically degraded broadcast channel, we can model any stochastically degraded broadcast channel as a physically degraded broadcast channel with the same marginal transition probabilities.

Theorem 1 ([2] [4]) The capacity region for the two-user degraded broadcast channel $X \rightarrow Y_1 \rightarrow Y_2$ is the convex hull of the closure of all (R_1, R_2) satisfying

$$R_2 \leq I(X_2; Y_2) \qquad R_1 \leq I(X; Y_1|X_2), \tag{1}$$

for some joint distribution $p(x_2)p(x|x_2)p(y, z|x)$, where the

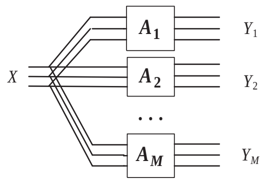

Fig. 2. Broadcast channel.

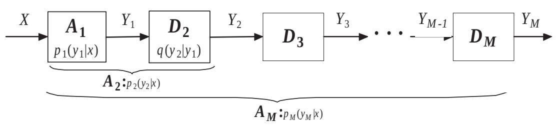

Fig. 3. Physically degraded broadcast channel.

auxiliary random variable X_2 has cardinality bounded by $|\mathcal{X}_2| \leq \min\{|\mathcal{X}|, |\mathcal{Y}_1|, |\mathcal{Y}_2|\}$.

B. The broadcast Z channel

The Z channel is a binary-asymmetric channel with $\Pr\{y = 0|x = 1\} = 0$ (see Fig. 1(a)). If symbol 1 is transmitted, symbol 1 is received with probability 1. If symbol 0 is transmitted, symbol 1 is received with probability α and symbol 0 is received with probability $1 - \alpha$. We can consider a Z channel as the OR operation of the channel input X and Bernoulli noise N with parameter α (see Fig. 4(a)). The diagram of a two-user broadcast Z channel is shown in Fig. 1(b). Because broadcast Z channels are stochastically degraded, we can model any broadcast Z channel as a physically degraded broadcast Z channel as shown in Fig. 4(b), where $\alpha_\Delta = \frac{\alpha_2 - \alpha_1}{1-\alpha_1}$ and $\alpha_2 \geq \alpha_1$.

III. Optimal Transmission Strategy for the Two-User Broadcast Z Channel

The communication system for the two-user broadcast Z channel is shown in Fig. 5. In a general scheme, the transmitter jointly encodes the independent messages W_1 and W_2. The receivers decode the noisy signals without collaboration. Since the broadcast Z channel is stochastically degraded, its capacity region can be found directly from Theorem 1. The capacity region for the broadcast Z channel $X \to Y_1 \to Y_2$ (see Fig. 6) is the convex hull of the closure of all (R_1, R_2) satisfying

$$\begin{aligned} R_2 \leq\ & I_2 = I(X_2; Y_2) \\ =\ & H\big((p_2\gamma + q_2 q_1)(1-\alpha_2)\big) - \\ & p_2 H\big(\gamma(1-\alpha_2)\big) - q_2 H\big(q_1(1-\alpha_2)\big), \end{aligned} \quad (2)$$

$$\begin{aligned} R_1 \leq\ & I_1 = I(X; Y_1|X_2) \\ =\ & p_2\big(H(\gamma(1-\alpha_1)) - \gamma H(1-\alpha_1)\big) + \\ & q_2\big(H(q_1(1-\alpha_1)) - q_1 H(1-\alpha_1)\big), \end{aligned} \quad (3)$$

for some probabilities p_1, p_2, γ, where $H(p)$ is the binary entropy function, $q_1 = 1 - p_1$, $q_2 = 1 - p_2$ and

$$\alpha_2 = \Pr\{y_2 = 1|x = 0\} = 1 - (1-\alpha_1)(1-\alpha_\Delta). \quad (4)$$

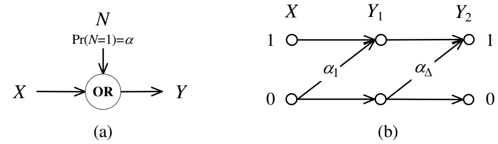

Fig. 4. (a) OR operation view of Z channel. (b) Physically degraded broadcast Z channel.

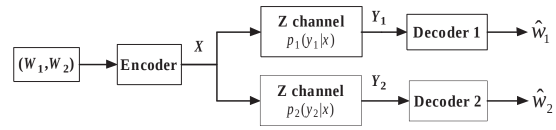

Fig. 5. Communication system for 2-user broadcast Z channel.

A. An independent encoding scheme

The previously defined broadcast communication system has a joint encoder which is potentially quite complex. We propose an independent encoding scheme for the broadcast Z channel, which can achieve the optimal surface of the capacity region. Fig. 7 shows the system diagram of the independent encoding scheme. First the messages W_1 and W_2 are encoded separately and independently. X_1 and X_2 are two binary random variables with $\Pr\{X_j = 1\} = p_j$ and $\Pr\{X_j = 0\} = q_j$. Thus $p_j + q_j = 1, j = 1, 2$. The transmitter broadcasts X, which is the OR of X_1 and X_2. We will show that by appropriately choosing the distribution of X_1 and X_2, i.e. q_1 and q_2, we can achieve any transmission rate pair in the optimal surface of the capacity region. The corresponding information theoretic diagram is Fig. 6 with $\gamma = 0$. Letting $\gamma = 0$ in equation (2) and (3), the achievable region for the broadcast Z channel $X \to Y_1 \to Y_2$, with the independent encoding scheme of Fig. 7, is the closure of all (R_1, R_2) satisfying

$$R_2 \leq I_2 = H(q_2 q_1(1-\alpha_2)) - q_2 H(q_1(1-\alpha_2)), \quad (5)$$

$$R_1 \leq I_1 = q_2 H(q_1(1-\alpha_1)) - q_2 q_1 H(1-\alpha_1), \quad (6)$$

for some $0 \leq q_1, q_2 \leq 1$.

Let us prove that any transmission rate pair in the optimal surface of the capacity region can be achieved using the independent encoding scheme of Fig. 7 with appropriate distributions of X_1 and X_2.

B. Optimal transmission strategy

Each particular choice of (p_1, p_2, γ) in Fig. 6 gives a particular transmission strategy and a rate pair (I_1, I_2). We

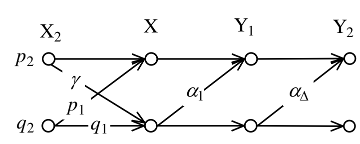

Fig. 6. Information theoretic diagram of the system.

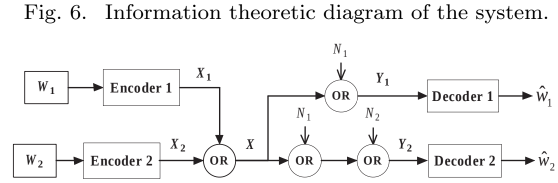

Fig. 7. Optimal transmission strategy for broadcast Z channels.

say that the optimal surface of a capacity region is the set of all Pareto optimal points (I_1, I_2), which are points for which it is impossible to increase rate I_1 without decreasing rate I_2 or vice versa. A transmission strategy is optimal if and only if it achieves a rate pair in the optimal surface.

Theorem 2: Every rate pair in the optimal surface of the capacity region for a broadcast Z channel with $0 < \alpha_1 < \alpha_2 < 1$ can be achieved with the independent encoding scheme shown in Fig. 7. In other words, every rate pair in the optimal surface of the capacity region can be achieved with $\gamma = 0$ in Fig. 6.

Before proving Theorem 2, we present and prove some preliminary results. From equations (2) and (3), we can see that the transmission strategies $(1-\gamma, 1-p_2, 1-p_1)$ and (p_1, p_2, γ) have the same transmission rate pairs. So we can assume $\gamma \le 1 - p_1$ in the rest of the section without loss of generality.

Theorem 3: For a broadcast Z channel with $0 < \alpha_1 < \alpha_2 < 1$, any transmission strategy (p_1, p_2, γ) with $0 < p_2 < 1, 0 < \gamma < 1 - p_1$ is not optimal.

The proof is given in Appendix A.

Corollary 1: The capacity region for the broadcast Z channel $X \to Y_1 \to Y_2$ is the convex hull of the achievable region with the independent encoding scheme.

Proof: From Theorem 3, we know that the transmission strategy (p_1, p_2, γ) is optimal only if at least one of these four equations $p_2 = 0$, $p_2 = 1$, $\gamma = 1 - p_1$, $\gamma = 0$ is true. When $p_2 = 0$, $p_2 = 1$ or $\gamma = 1 - p_1$, the transmission rate for the second user, I_2 in equation (2), is zero, which means that the only optimal rate pair that can be achieved is point B in Fig. 13. Point B can also be achieved by the independent encoding scheme with $\gamma = 0$, $p_2 = 0$ and $p_1 = \arg\max(H((1-x)(1-\alpha_1)) - (1-x)H(1-\alpha_1))$. Thus, all the optimal rate pairs in the optimal surface of the capacity region can be achieved by using the independent encoding scheme with $\gamma = 0$ and time sharing.

The achievable region given by equations (5) and (6) is not in an explicit form. In order to find the explicit form of the achievable region, we consider the following optimization problem: maximize $\lambda I_1 + (1-\lambda) I_2$ for any fixed $\lambda \in [0, 1]$.

Theorem 4: The optimal solution to the maximization problem

$$\begin{aligned} \text{maximize} \quad & \lambda I_1 + (1-\lambda) I_2 \\ \text{subject to} \quad & I_2 = H(q_2 q_1 (1-\alpha_2)) - q_2 H(q_1(1-\alpha_2)) \\ & I_1 = q_2 H(q_1(1-\alpha_1)) - q_2 q_1 H(1-\alpha_1) \\ & 0 \le q_2 \le 1, 0 \le q_1 \le 1, \end{aligned} \tag{7}$$

is unique and it is given below for any fixed $\lambda \in [0, 1]$. Define

$$\varphi(x) = \frac{\log(1-(1-\alpha_2)x)}{\log(1-(1-\alpha_1)x) + \log(1-(1-\alpha_2)x)}, \tag{8}$$

$$\psi(x) = \frac{1}{x e^{H(x)/x} + x}. \tag{9}$$

Case 1: if $\varphi(\psi(1-\alpha_1)) \le \lambda \le 1$, then the optimal solution is $q_2^* = 1, q_1^* = \psi(1-\alpha_1)$ and the corresponding rate pair is $I_1^* = H(q_1^*(1-\alpha_1)) - q_1^* H(1-\alpha_1), I_2^* = 0$.

Case 2: if $0 \le \lambda \le \varphi(1)$, then the optimal solution is $q_2^* = \psi(1-\alpha_2), q_1^* = 1$ and the corresponding rate pair is $I_1^* = 0, I_2^* = H(q_2^*(1-\alpha_2)) - q_2^* H(1-\alpha_2)$.

Case 3: if $\varphi(1) < \lambda < \varphi(\psi(1-\alpha_1))$, then the optimal solution has

$$\lambda \log(1 - q_1^*(1-\alpha_1)) = (1-\lambda)\log(1 - q_1^*(1-\alpha_2)), \tag{10}$$

$$H(q_1^*(1-\alpha_2)) - q_1^*(1-\alpha_2)\log\frac{1-q_2^* q_1^*(1-\alpha_2)}{q_2^* q_1^*(1-\alpha_2)} = \frac{\log(1-q_1^*(1-\alpha_2))}{\log(1-q_1^*(1-\alpha_1))} \cdot \big(H(q_1^*(1-\alpha_1)) - q_1^* H(1-\alpha_1)\big). \tag{11}$$

The proof is given in Appendix B. The achievable region is shown with two upper bounds in Fig. 8(a). From case 1, we can see that point A corresponds to the largest transmission rate for the second user. The first upper bound is the tangent of the achievable region in point A, and its slope is $-\varphi(1)/(1-\varphi(1))$. From case 2, we show that point B provides the largest transmission rate for the first user. The second upper bound is the tangent of the achievable region in point B, and its slope is $-\varphi(\psi(1-\alpha_1))/(1-\varphi(\psi(1-\alpha_1)))$. Case 3 gives us the optimal surface of the achievable region except points A and B.

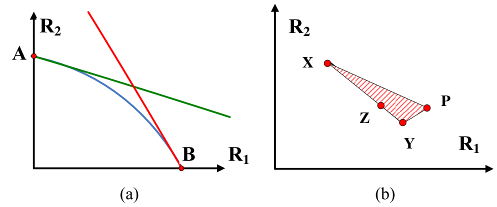

Fig. 8. (a) The capacity region and two upper bounds. (b) Point Z can not be on the boundary of the capacity region.

Given α_1 and α_2, which completely describe a two-user degraded broadcast Z channel, the capacity region can be explicitly described. Cases 1 and 2 identify the corner points of the capacity region. From Cases 1, 2, and 3 above, the rest of the curve is described by the following range of q_1 values:

$$\psi(1-\alpha_1) < q_1 < 1. \tag{12}$$

The associated q_2 values follow from (11). The curve of the capacity region boundary follows from using the q_1 and q_2 values in (5) and (6). For example, for $\alpha_1 = 0.15$ and $\alpha_2 = 0.6$, the range of q_1 values is $0.445 < q_1 < 1$ and the associated capacity region boundary is plotted in Fig. 12.

Finally, we prove Theorem 2.

Proof by contradiction: Suppose the point Z in Fig. 8(b) is on the boundary of the capacity region for the broadcast Z channel but not in the achievable region with the independent encoding scheme. Thus, it can be achieved only by time sharing of points X and Y, which is in the achievable region. Clearly, The slope of the line segment XY is neither zero nor infinity. Suppose the slope of XY is $-k, 0 < k < \infty$, so points X and Y provide the same $k \cdot R_1 + R_2$. From Theorem 4, the optimal solution to the maximization problem $\max(\lambda I_1 + (1-\lambda) I_2))$ is unique, therefore neither X nor Y maximizes

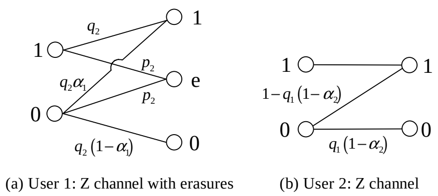

Fig. 9. Perceived channel by each decoder.

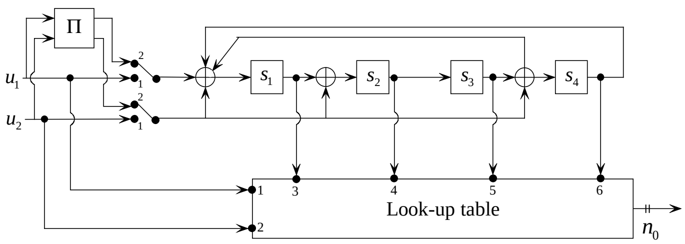

Fig. 10. 16-state nonlinear turbo code structure, with $k_0 = 2$ input bits per trellis section.

$(k \cdot I_1 + I_2)$. Thus, there exists an achievable point P on the right upper side of the line XY and the triangle $\triangle XYP$ is in the capacity region. So the point Z must not be on the boundary of the capacity region (contradiction).□

IV. Nonlinear-Turbo Codes for the Two-User Broadcast Z Channel

In this section we show a practical implementation of the transmission strategy for the two-user broadcast Z channel. As proved in Section III, the optimal surface is achieved by transmitting the OR of the encoded data of each user, provided that the density of ones of each of these encoded streams is chosen properly. Hence, a family of codes that provides a controlled density of ones and zeros is required. We propose the use of nonlinear turbo codes, introduced in [6]. Nonlinear turbo codes are parallel concatenated trellis codes, with k_0 input bits and n_0 output bits per trellis section. A look-up table assigns the output label of each branch of the trellis, so that the required ones density is achieved. Each constituent encoder for the turbo code in this paper is a 16-state trellis code, with $k_0 = 2$, with trellis structure shown in Fig. 10. The output labels are assigned via a constrained search that provides the required ones density for each case, using the tools presented in [6] for the Z Channel. Also, the tools presented in [6] were general enough to be applied to the Z Channel with erasures as perceived by user 1.

Receiver 1 uses sequential decoding as shown in Fig. 11. Denote as $\hat{X}_2$ the decoded stream corresponding to user 2. Since the transmitted data is $x = x_1(\mathrm{OR})x_2$, whenever a bit $x_2 = 1$, there is no information about x_1, and x_1 can be considered an erasure. Hence, the input stream to Decoder 1 is

$$\hat{y}_1 = e(y_1, \hat{x}_2) = \begin{cases} y_1 & \text{if} \quad \hat{x}_2 = 0, \\ e & \text{if} \quad \hat{x}_2 = 1. \end{cases} \tag{13}$$

Therefore, Decoder 2 sees a Z Channel with erasures as shown in Fig. 9. Note that if α_1 is much smaller than α_2 we can use hard decoding in Decoder 2 instead of soft decoding without any loss in performance. Since the code for user 2 is designed for a Z Channel with 0-to-1 crossover probability $1-(1-\alpha_2)q_1$, and the channel perceived by Decoder 2 in user 1 is a Z-Channel with crossover probability $1-(1-\alpha_1)q_1 < 1-(1-\alpha_2)q_1$, the bit error rate of $\hat{x}_2$ is negligible compared to the bit error rate of Decoder 1. In fact, in all the simulations shown in Section V, which include 100 frame errors of user 1, none of the errors were produced by Decoder 2.

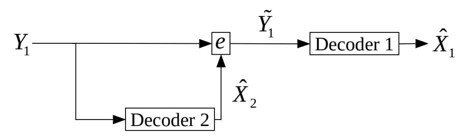

Fig. 11. Decoder structure for user 1.

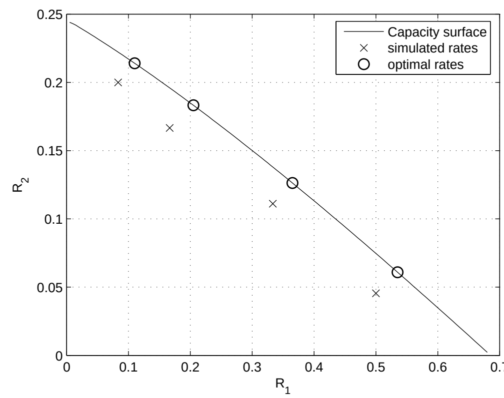

Fig. 12. Broadcast Z channel with crossover probabilities $\alpha_1 = 0.15$ and $\alpha_2 = 0.6$ for receiver 1 and 2 respectively: achievable capacity region, simulated rate pairs (R_1, R_2) and their corresponding optimal rates.

V. Results

We have simulated the transmission strategy for the two-user broadcast Z channel with crossover probabilities $\alpha_1 = 0.15$ and $\alpha_2 = 0.6$, using nonlinear turbo codes, with the structure shown in Fig. 10. Fig. 12 shows the achievable region of the rate pairs (R_1, R_2) on this channel, and the simulated rate pairs. It also shows the optimal rate pairs used to compute the ones densities of each code. For each of these four simulated rate pairs, the loss in mutual information from the associated optimal rate is only 0.04 bits or less in R_1 and only 0.02 bits or less in R_2. Table I shows bit error rates for each rate pair, the ones densities p_1 and p_2, and the interleaver lengths K_1 and K_2 used for each code. For simplicity, we chose K_1 and K_2 so that the codeword length n would be the same for user 1 and user 2, except for rate pairs $R_1 = 1/2$ and $R_2 = 1/22$, where one codeword length of user 2 is twice the length of user 1.

VI. Conclusions

This paper presented an optimal transmission strategy for the broadcast Z channel with simple encoding and decoding. We proved that any point in the optimal surface of the capacity region can be achieved by independently encoding the messages corresponding to different users and

TABLE I

BER for two-user broadcast Z channel with crossover probabilities $\alpha_1 = 0.15$ and $\alpha_2 = 0.6$.

R_1	R_2	p_1	p_2	K_1	K_2	BER_1	BER_2
1/12	1/5	0.106	0.56	4800	1700	2.54×10^{-5}	1.24×10^{-5}
1/6	1/6	0.196	0.5	2048	2048	7.01×10^{-6}	5.33×10^{-6}
1/3	1/9	0.336	0.3739	4608	1536	7.13×10^{-6}	6.70×10^{-6}
1/2	1/22	0.463	0.1979	5632	1024	9.27×10^{-7}	3.27×10^{-6}

transmitting the OR of the encoded signals. Also, the distributions of the outputs of each encoder that achieve the optimal surface were provided. Nonlinear-turbo codes that provide a controlled distribution of ones and zeros in the codes were used to demonstrate a low-complexity scheme that works close to the optimal surface.

Appendices

Appendix A

Here we prove Theorem 3. In (2) and (3), denote

$$I_1(p_1,p_2,\gamma) = I(X;Y_1|X_2)\big|_{p_1,p_2,\gamma} \quad (14)$$

$$I_2(p_1,p_2,\gamma) = I(X_2;Y_2)\big|_{p_1,p_2,\gamma} \quad (15)$$

$$I_{1,2}(p_1,p_2,\gamma) = (I_1,I_2)\big|_{p_1,p_2,\gamma}. \quad (16)$$

The strategy (p_1,p_2,γ) has the rate pair $I_{1,2}(p_1,p_2,\gamma)$. The theorem is true if we can increase both I_1 and I_2 when $0<p_2<1, 0<\gamma<1-p_1$.

Firstly we compare the strategies (p_1,p_2,γ) and $(p_1-p_2\delta_1,p_2,\gamma-q_2\delta_1)$ for a small positive number δ_1.

$$\begin{aligned}\Delta_1 I_j &= I_j(p_1-p_2\delta_1,p_2,\gamma-q_2\delta_1)-I_j(p_1,p_2,\gamma)\\ &\simeq \frac{\partial I_j(p_1-p_2\delta_1,p_2,\gamma-q_2\delta_1)}{\partial\delta_1}\Big|_{\delta_1=0}\delta_1\\ &= (-1)^j q_2p_2(1-\alpha_j)\big(\log\frac{1-\gamma(1-\alpha_j)}{\gamma(1-\alpha_j)}\\ &\quad +\log\frac{q_1(1-\alpha_j)}{1-q_1(1-\alpha_j)}\big)\delta_1, j=1,2. \quad (17)\end{aligned}$$

The small change of the rate pair $(\Delta_1 I_1, \Delta_1 I_2)$ is shown Fig. 13. Point A is the rate pair of the strategy (p_1,p_2,γ), the arrow Δ_1 shows the small movement of the rate pair $(\Delta_1 I_1, \Delta_1 I_2)$.

Secondly we compare the strategies (p_1,p_2,γ) and $(p_1-(\gamma-q_1)\delta_2, p_2-q_2\delta_2,\gamma)$ for a small positive number δ_2.

$$\begin{aligned}\Delta_2 I_j &= I_j(p_1-(\gamma-q_1)\delta_2,p_2-q_2\delta_2,\gamma)\\ &\quad -I_j(p_1,p_2,\gamma)\\ &\simeq \frac{\partial I_j(p_1-(\gamma-q_1)\delta_2,p_2-q_2\delta_2,\gamma)}{\partial\delta_2}\Big|_{\delta_2=0}\delta_2\\ &= (-1)^j q_2\delta_2\big\{\gamma(1-\alpha_j)\log\frac{q_1}{\gamma}+\\ &\quad (1-\gamma(1-\alpha_j))\log\frac{1-q_1(1-\alpha_j)}{1-\gamma(1-\alpha_j)}\big\}\\ &= (-1)^j q_2\delta_2 D(\gamma(1-\alpha_j)\parallel q_1(1-\alpha_j)),\\ &\quad j=1,2. \quad (18)\end{aligned}$$

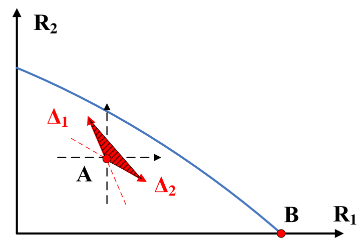

Fig. 13. Capacity region and the changes of rate pairs.

where $D(p \parallel q)$ is the relative entropy between distribution p and q. The arrow Δ_2 in Fig. 13 shows the small movement of the rate pair $(\Delta_2 I_1, \Delta_2 I_2)$.

Now we show that

$$\frac{\Delta_1 I_2}{\Delta_1 I_1} < \frac{\Delta_2 I_2}{\Delta_2 I_1} < 0. \quad (19)$$

$$\begin{aligned} & \frac{\Delta_1 I_2}{\Delta_1 I_1} < \frac{\Delta_2 I_2}{\Delta_2 I_1}\\ \Leftrightarrow\ & \frac{D(\gamma(1-\alpha_2)\|q_1(1-\alpha_2))+\log\frac{1-\gamma(1-\alpha_2)}{1-q_1(1-\alpha_2)}}{D(\gamma(1-\alpha_1)\|q_1(1-\alpha_1))+\log\frac{1-\gamma(1-\alpha_1)}{1-q_1(1-\alpha_1)}}\\ & > \frac{D(\gamma(1-\alpha_2)\|q_1(1-\alpha_2))}{D(\gamma(1-\alpha_1)\|q_1(1-\alpha_1))}\\ \Leftrightarrow\ & \frac{D(\gamma(1-\alpha_1)\|q_1(1-\alpha_1))}{\log\frac{1-\gamma(1-\alpha_1)}{1-q_1(1-\alpha_1)}} > \frac{D(\gamma(1-\alpha_2)\|q_1(1-\alpha_2))}{\log\frac{1-\gamma(1-\alpha_2)}{1-q_1(1-\alpha_2)}}\\ \Leftrightarrow\ & f(x) = \frac{D(\gamma x\|q_1 x)}{\log\frac{1-\gamma x}{1-q_1 x}} \text{is monotonically increasing}\\ & \text{in } \{x|0<x<1\}\\ \Leftrightarrow\ & f'(x) = \big(\log\frac{\gamma x}{q_1 x}\log\frac{1-\gamma x}{1-q_1 x} - (\log\frac{1-\gamma x}{1-q_1 x})^2\\ & +\log\frac{\gamma x}{q_1 x}(\frac{1}{1-\gamma x}-\frac{1}{1-q_1 x})\big)\gamma(\log\frac{1-\gamma x}{1-q_1 x})^{-2} > 0. \quad (20)\end{aligned}$$

Let $u = 1-\gamma x$ and $v = 1-q_1 x$. So we have $0<v<u<1$ and need to prove that

$$g(u,v) = \log\frac{u}{v}\log\frac{1-u}{1-v} - (\log\frac{u}{v})^2 + \log\frac{1-u}{1-v}(\frac{1}{u}-\frac{1}{v}) > 0. \quad (21)$$

Since $g(v,v) = 0, \forall 0<v<1$, we just need to show that $\frac{\partial g(u,v)}{\partial u} > 0 \quad \forall 0<v<u<1$. For a fixed u, we can consider $\phi_u(v) = \frac{\partial g(u,v)}{\partial u}$ as a function of v. Because $\phi_u(u) = \frac{\partial g(u,v)}{\partial u}\big|_{v=u} = 0, \forall 0<u<1$, we only need to prove that $\frac{\partial^2 g(u,v)}{\partial u\partial v} < 0$. It is easy to check that $\forall 0<v<u<1$

$$\frac{\partial^2 g(u,v)}{\partial u\partial v} = -\frac{(u-v)^2}{u^2v^2(1-u)(1-v)} < 0. \quad (22)$$

Thus, the inequality (19) is true, which means that the slope of Δ_1 is smaller than that of Δ_2 in Fig. 13. The achievable shaded region is on the upper right side of point A. Therefore, we can increase the rate pair $I_{1,2}(p_2, \gamma, p_1)$ together and the strategy (p_2, γ, p_1) is not optimal.$\square$

Appendix B

Here we prove Theorem 4. In problem (7), the objective function $\lambda I_1 + (1-\lambda) I_2$ is bounded and the domain $0 \le q_1, q_2 \le 1$ is closed, so the maximum exists and can be attained. First we discuss some possible optimal solutions and then we show that only one of them is the optimum for any fixed λ between 0 and 1.

Case 0: If $q_1 = 0$ or $q_2 = 0$ or $q_1 = q_2 = 1$, then $I_1 = I_2 = 0$ and so it can not be the optimum.

Case 1: If $q_2 = 1$ and $0 < q_1 < 1$, then $I_2 = 0$ and

$$\frac{\partial I_1}{\partial q_1} = (1-\alpha_1) \log \frac{1 - q_1(1-\alpha_1)}{q_1(1-\alpha_1)} - H(1-\alpha_1) = 0 \quad (23)$$

$$\Rightarrow q_1^* = \frac{1}{(1-\alpha_1)(e^{H(1-\alpha_1)/(1-\alpha_1)} + 1)}. \quad (24)$$

Case 2: If $q_1 = 1$ and $0 < q_2 < 1$, then $I_1 = 0$ and

$$\frac{\partial I_2}{\partial q_2} = (1-\alpha_2) \log \frac{1 - q_2(1-\alpha_2)}{q_2(1-\alpha_2)} - H(1-\alpha_2) = 0 \quad (25)$$

$$\Rightarrow q_2^* = \frac{1}{(1-\alpha_2)(e^{H(1-\alpha_2)/(1-\alpha_2)} + 1)}. \quad (26)$$

Case 3: If $0 < q_1, q_2 < 1$, then the optimum is attained when

$$q_2 \frac{\partial(\lambda I_1 + (1-\lambda) I_2)}{\partial q_2} + q_1 \frac{\partial(\lambda I_1 + (1-\lambda) I_2)}{\partial q_1} = 0$$

$$\Rightarrow \lambda \log(1 - q_1^*(1-\alpha_1)) = (1-\lambda) \log(1 - q_1^*(1-\alpha_2)), \quad (27)$$

and

$$\begin{aligned} &\tfrac{\partial(\lambda I_1 + (1-\lambda) I_2)}{\partial q_2} = 0 \\ &\Rightarrow (1-\lambda)\big(H(q_1^*(1-\alpha_2)) - q_1^*(1-\alpha_2) \log \tfrac{1 - q_2^* q_1^*(1-\alpha_2)}{q_2^* q_1^*(1-\alpha_2)}\big) \\ &\quad -\lambda\big(H(q_1^*(1-\alpha_1)) - q_1^* H(1-\alpha_1)\big) = 0 \\ &\Rightarrow H(q_1^*(1-\alpha_2)) - q_1^*(1-\alpha_2) \log \tfrac{1 - q_2^* q_1^*(1-\alpha_2)}{q_2^* q_1^*(1-\alpha_2)} = \\ &\tfrac{\log(1 - q_1^*(1-\alpha_2))}{\log(1 - q_1^*(1-\alpha_1))} \cdot \big(H(q_1^*(1-\alpha_1)) - q_1^* H(1-\alpha_1)\big). \quad (28) \end{aligned}$$

Now we are going to find which case is optimal for different λ.

Lemma 1: Function $l(q_1) = \frac{\log(1 - q_1(1-\alpha_2))}{\log(1 - q_1(1-\alpha_1))}$ is monotonically decreasing in the domain $0 \le q_1 \le 1$ when $\alpha_1 < \alpha_2$

Lemma 2: The solution to equation (27) exists in $(0, 1)$ and is unique for any λ with $\varphi(1) < \lambda < \varphi(0)$.

Proof: Equation (27) is equivalent to $l(q_1^*) = \lambda/(1-\lambda)$. From Lemma 1, $l(q_1)$ is monotonically decreasing. Therefore, when $l(1) < \lambda/(1-\lambda) < l(0)$, i.e. $\varphi(1) < \lambda < \varphi(0)$, the solution $q-1^*$is unique and $q_1^* \in (0, 1)$.$\square$

Lemma 3: The unique solution (q_1^*, q_2^*) to equation (27) and (28) in case 3 is the optimum if $\varphi(1) < \lambda < \varphi(\psi(1-\alpha_1))$.

Proof: From Lemma 2, the solution q_1^* to equation (27) is unique if $\varphi(1) < \lambda < \varphi(\psi(1-\alpha_1))$. From (28)

$$\begin{aligned} m(q_2) = &\big(H(q_1^*(1-\alpha_2)) - q_1^*(1-\alpha_2) \log \tfrac{1 - q_2 q_1^*(1-\alpha_2)}{q_2 q_1^*(1-\alpha_2)}\big) * \\ &\log(1 - q_1^*(1-\alpha_1)) - \big(H(q_1^*(1-\alpha_1)) - q_1^* H(1-\alpha_1)\big) * \\ &\log(1 - q_1^*(1-\alpha_2)) = 0. \end{aligned} \quad (29)$$

Clearly, $m(q_2)$ is monotonically increasing and

$$\lim_{q_2 \to 0} m(q_2) = -\infty < 0. \quad (30)$$

$$\begin{aligned} &\varphi(1) < \lambda < \varphi(\psi(1-\alpha_1)) \\ \Rightarrow \quad &q_1^* > \psi(1-\alpha_1) \\ \Rightarrow \quad &m(1) > 0. \end{aligned} \quad (31)$$

That means the unique solution q_2^* to equation (28) is in the domain $0 \le q_2 \le 1$. Furthermore, when $\varphi(1) < \lambda < \varphi(\psi(1-\alpha_1))$, case 1 and case 2 can not be the optimum because

$$\frac{\partial(\lambda I_1 + (1-\lambda) I_2)}{\partial q_2}\Big|_{q_2=1, q_1=\psi(1-\alpha_1)} < 0, \quad (32)$$

$$\frac{\partial(\lambda I_1 + (1-\lambda) I_2)}{\partial q_1}\Big|_{q_1=1, q_2=\psi(1-\alpha_2)} < 0. \quad (33)$$

Therefore, case 3 is the optimum.$\square$

Lemma 4: The unique solution $(q_2^* = 1, q_1^* = \psi(1-\alpha_1))$ in case 1 is the optimum if $\varphi(\psi(1-\alpha_1)) \le \lambda \le 1$.

Proof: When $\varphi(\psi(1-\alpha_1)) \le \lambda \le 1$, case 3 is not optimal because there is no solution $q_1 \in (0, 1)$ to equation (27). Case 2 is not optimal because inequality (33) holds. So case 1 is the optimum.$\square$

Lemma 5: The unique solution $(q_2^* = \psi(1-\alpha_2), q_1^* = 1)$ in case 2 is the optimum if $0 \le \lambda \le \varphi(1)$.

Proof: When $0 \le \lambda \le \varphi(1)$, case 3 is not optimal because there is no solution $q_2 \in (0, 1)$ to equation (28). Case 1 is not optimal because inequality (32) holds. So case 2 is the optimum.$\square$

From Lemma 3,4 and 5, Theorem 4 is immediately proved.$\square$

References

[1] T. M. Cover. Broadcast channels. *IEEE Trans. Inform. Theory*, IT-18:2–14, January 1972.

[2] P. P. Bergmans. Random coding theorem for broadcast channels with degraded components. *IEEE Trans. Inform. Theory*, IT-19:197–207, March 1973.

[3] P. P. Bergmans. A simple converse for broadcast channels with additive white Gaussian noise. *IEEE Trans. Inform. Theory*, IT-20:279–280, March 1974.

[4] R. G. Gallager. Capacity and coding for degraded broadcast channels. *Probl. Pered. Inform.*, 10:3–14, July–Sept. 1974.

[5] T. M. Cover. Comments on Broadcast channels. *IEEE Trans. Inform. Theory*, 44:2524–2530, October 1998.

[6] M. Griot, A.I. Vila Casado, and R.D. Wesel. Non-linear turbo codes for interleaver-division multiple access on the OR channel. In *GLOBECOM '06. IEEE Global Telecomm. Conf.*, 27 Nov. - 1 Dec. 2006.